

Submitter Name: Hélène Choquet
Submitted email: Helene.Choquet@kp.org

Genetic Ancestry and Tobacco Smoking Behavior

Hélène Choquet¹, Jie Yin¹, Thomas J. Hoffmann^{2,3}, Yambazi Banda², Mark N. Kvale², Catherine Schaefer¹, Neil Risch^{1,2,3}, Constance Weisner^{1,4}, and Eric Jorgenson¹

¹Division of Research, Kaiser Permanente Northern California (KPNC); ²Institute for Human Genetics, University of California San Francisco; ³Department of Epidemiology and Biostatistics, UCSF; ⁴Department of Psychiatry and Weill Institute for Neurosciences, UCSF

Tobacco smoking contributes to numerous common diseases and the combined attributable risk makes it one of the leading causes of death in the United States. Smoking behavior varies widely across ethnic groups, but it is not clear why. Here, we examined the contribution of genetic ancestry in tobacco smoking initiation in 41,924 smokers (former/current) and 56,609 non-smokers (never) from the Kaiser Permanente Northern California Genetic Epidemiology Research in Adult Health and Aging (GERA) cohort. We examined the association of ancestry principal components (PCs) with smoking initiation within each GERA ethnic group (non-Hispanic white, Hispanic/Latino, East Asian, and African-American), adjusting for age and sex. Within each ethnic group, we observed significant associations between genetic ancestry and tobacco smoking initiation, except for African-Americans. In East-Asians, the first two PCs, which differentiate geographical clines across East Asia, were both strongly associated with smoking initiation ($P_{PC1}=2.3 \times 10^{-19}$; $P_{PC2}=1.3 \times 10^{-7}$), with the highest prevalence observed in northern East Asia (i.e., Japan). In non-Hispanic whites, individuals of northern and western European ancestry had higher prevalence of smoking ($P_{PC1}=0.03$; $P_{PC2}=2.9 \times 10^{-4}$) compared with individuals of southern and eastern European ancestry. In Hispanic/Latinos, we observed higher smoking prevalence with greater European (versus Native American) ancestry ($P_{PC1}=9.79 \times 10^{-14}$). We then ascertained whether ancestry associations were explained by 365 smoking initiation-associated genetic variants recently identified by the GWAS and Sequencing Consortium of Alcohol and Nicotine use. The genetic ancestry associations were not attenuated after including these genetic variants, suggesting that factors that correlate with genetic ancestry are more likely to be cultural and/or socioeconomic.